

Linguo-Methodical Objectives and Frequency-Specific Approaches for Kazakh Language Instruction

Aliya Kuryshzhanova
aliya.kuryshzhan@gmail.com
Hankuk University of Foreign Studies

Author's Note

This work was supported by the Hankuk University of Foreign Studies Research Fund.

Correspondence concerning this article should be addressed to:

Aliya Kuryshzhanova
Department of Central Asian Studies,
HUFs, Global campus, Wangsan-ri, Mohyeon-myeon, Cheoin-gu,
Yongin-si, Kyeonggi-do, 449-791, Korea

Abstract

Frequency-specific lexical lists provide objective criteria on which to base accurate level-specific content distribution, quantitative characteristics, adequate dosing and distribution of teaching materials by levels, objective assessment and others. This technique also makes the teaching objectives more concrete and visible. The modern Kazakh language and its contemporary vocabulary are reflected above all in its spoken form. Therefore, a beginning frequency list has been conducted of current Kazakh speech based on types of modern communication. In particular, we analyzed spoken conversations in movies (25000 units), mass media interviews and dialogues from literature essays (30000 units). Selected frequency-specific vocabulary helps to differentiate content for various educational levels differentiation, identifying nuclear / basic high frequency vocabulary for specific area and also identified objects for accurate assessment. Use of frequent vocabulary and grammar minimums allows reducing and detailing the studied content to what is actually needed within the usually very limited time of language courses. This way provides possibilities to solve multiple needs of Kazakh language researchers, methodologists and practice teachers.

Keywords: teaching methods, Kazakh language, frequency list

I. Introduction

Demographic growth of the Kazakh ethnic group, major changes in the community, and urban globalization have significantly increased the spheres of use of the state language. This in turn has caused major changes in the language's vocabulary. New expanded functions have to be created by use of the state language in such include services KazNet.kz, Uikibilim.kz, Google.kz, e-gov.kz, text semantical analyzers and SEO formalizers, search machines / searchers, and many others. Kazakhstan's language policy has stimulated the use of the state language by both native and other speakers, thus demanding a search for new corpus and status development of the language and new methods of instruction.

State policy seeks to achieve mastery of Kazakh by all citizens; full-fledged social functioning of the language, production of linguistic information and communication technologies, and many other goals. Al-

though state language policy identifies approaches to achieve many current linguistic objectives, problems of studying the state language will remain unfulfilled until the country attains the state of socially-comfortable poly-lingualism.

Although much research has been conducted about Kazakh linguistics, greater demand for «Kazakh for foreign / Russian-speakers» has become manifest since independence. Kazakh language methodology reviews the training problems from the following perspectives: a) Kazakh as L1KZ (first / native language, i.e. Kazakh for Kazakh-speaking) and b) Kazakh as L2KZ (second / non-native / foreign – for those speaking other language). The current language situation makes L2KZ particularly relevant, and this issue requires specific approaches to language acquisition.

The increased attention to methodological issues requires major substantiation of teaching parameters that affect the result: among these are detailed objectives of training, level-specific training and specialization, as well as such parameters as language environment expanding.

This means that it is necessary to various different methodologies a) when Russian-speaking Kazakhstan citizens study the state language as L2KZ in the language environment from the methodology required b) when Kazakh language is studied as L2KZ outside the language environment. For instance, specific methodologies are required for Kazakh language courses in Kazakh communities, for language training of foreign companies' employees, and for students of foreign universities. When preparing training courses and selecting methodologies, identification and using focused in frequency parameters of language leads towards more efficient performance results.

In order to find an efficient methodology, it is necessary to identify the objective criteria for a course, an accurate level-specific definition of the training content, quantitative counting and then the statistically distributed language material by study levels, objective assessment, this is best accomplished by applying frequency-specific language characteristics (frequency of the most often used words). Hierarchic frequency-specific lexical lists of educational language materials make the training objectives more concrete. However, existing Kazakh frequency vocabularies were created based on written texts, and they are outdated, since they were compiled in the last quarter of the last century. Meanwhile, modern Kazakh, with its more up-to-date vocabulary reflects the current state of the language. This first reflected in the verbal form, because verbal use more promptly reacts to changes and more adequately reflects real functioning of the current language.

The research for this paper has been conducted in order to analyze the current Kazakh language based on movie materials, mass media interviews and dialogues of Kazakh writers' literature. Data have been processed with specific software codes to formalize and create an algorithm of the language, and form ranked vocabulary databases for the verbal Kazakh language.

The research for this theme has been conducted by processing modern Kazakh speech and identifying the most frequently used vocabulary; these items were used to develop a primary education course. In selecting educational materials based upon frequency parameters it becomes obvious that various given characteristics of the Kazakh language significantly affect developing frequency vocabulary.

Let us compare frequency lists made primarily for modern Kazakh oral speech with the frequency vocabulary of Abai Kunanbayev's written texts. Significant differences of parameters of the applied vocabulary become obvious in comparing these two frequency lists. In this case the given parameters also define the nature of this text frequency vocabulary: written or spoken language. Tables 1 and 2 demonstrate the use of the first 20 most high-frequency lexical units, which differ by oral / written speech parameters.

The first table of these statistic materials are a part of our research for creating the frequency list of oral communication, which was generated by modern Kazakh movies spoken dialogues, and the second table is a part of frequency list of Dictionary of Abay Kunanbaev, located at Kazakh language portal (<http://til.gov.kz/wps/portal>).

Table 1 Modern Kazakh movies dialogues vocabulary

#	lexeme	Quant.	Frequency	Freq. %
1.	ал / ал(у)	662	0,265	26,48
2.	ай	602	0,241	24,08
3.	не	484	0,194	19,36
4.	ол	420	0,168	16,80
5.	бол(у)	370	0,148	14,80
6.	ғой	267	0,107	10,68
7.	бір	250	0,100	10,00
8.	жоқ	245	0,098	9,80
9.	де / де(у)	228	0,091	9,12
10.	ой	223	0,089	8,92
11.	ма	218	0,087	8,72
12.	бар / бар(у)	211	0,084	8,44
13.	сен / сен(у)	203	0,081	8,12
14.	қыз	200	0,080	8,00
15.	ақ	194	0,078	7,76
16.	өз	185	0,074	7,40
17.	кел(у)	182	0,073	7,28
18.	ба	177	0,071	7,08
19.	он	175	0,070	7,00
20.	мен	167	0,067	6,68

Table 2 Dictionary of Abay Kunanbayev

#	Frequency	lexeme
1.	69	болу
2.	56	айту
3.	56	алу
4.	50	көру
5.	45	білу
6.	44	келу
7.	40	қалу
8.	39	беру
9.	39	кету
10.	35	түру
11.	34	жүру
12.	32	деу
13.	25	өз
14.	23	жату
15.	23	сөз
16.	23	кою
17.	22	есту
18.	22	жету
19.	22	отыру
20.	22	салу

In this kind of collation the following characteristic becomes visible this distinction and it proves the most important in this situation: this vocabulary use oral or written speech. These two tables of the most often used 20 words show that spoken and written Kazakh use different vocabulary or even there are the same words they are used with different range of frequency. Besides, other important parameters are: obsolete vocabulary or current neologisms; topics chosen, style and text content as well as some other specific ones.

It is worth mentioning that because these findings demonstrate frequency analysis findings, those conducted at various period, based on various methodologies and with various technologies (e.g. analysis of the movies language became possible due to new technology methods), then parameters of these lists have different characteristics. Thus, a frequency list of Kazakh movies demonstrate more detailed and disseminated frequency characteristics, which provide more information for further use. Although such comparison is not absolutely precise, this comparison of various type frequency vocabularies is obvious and provides more detailed materials for practitioners. Besides, previous years' data are valuable in serving the object of synchronic and diachronic analysis.

The case illustrated above shows the difference between the oral and written forms of the language. The lists for both contain similar lexeme units of the most frequent vocabulary, which falls into top 20; however, these lexical units indicate various frequency ranges, thus characterizing their various application in oral and written communication.

Awareness of this information is important for many practical applied objectives, one of them is the teaching Kazakh. In the present case we have looked within the focus of educational purposes, where definition of lexical minimum for levels of education, their adequate control and evaluation are always critical. If teachers possess such information as what vocabulary is important for a particular level; what specific language is re-

quired for a particular profession; or what vocabulary is important for particular audience or goals – they obtain more opportunities to develop more targeted and time-efficient educational course, which is critical within the limited length of an educational course.

Besides, frequency lists based on written speech texts and oral speech, obviously indicate what has to be included into the oral communication course, which vocabulary has to be emphasized, e.g. within the academic writing course – differences of frequency vocabularies of oral and written speech are obvious enough. Therefore, besides the obviously most nuclear vocabulary, the specific objective of teaching – oral communication or writing – should be considered while developing every language course.

It is also worth mentioning the widely disseminated practice is that the teachers emphasize semantically important vocabulary for the topic of a lesson or course. This is usually justified with a significant rule of thumb. However, frequency analysis findings indicate that semantically zero or less significant, or semantically insignificant connective words are the most applicable. For language teachers this means that equal attention should be paid to teaching this particular vocabulary too. Work on structurally formalized and abstract models of statements, usually built upon most frequent connective words, is a good example of this: e.g. if the students have been taught to understand and work using the «Kim ony zhasajdy?» / «Neni kalaj zhasajdy?» / «Kim ony kalaj zhasajdy?» and other models, then the following semantically lexical replenishment and model variation gets mastered by them more successfully.

Let us now compare frequency lists for Kazakh language (the analysis gives data for current Kazakh mass media language – Table 3) and, for instance, frequency educational dictionary of Russian language (N. J. Brown. Russian Learners' dictionary 10000 words in frequency order. 2003 – Table 4). Both the first and second examples demonstrate high ranges of frequency of helping words in particular:

#	lexeme	Frequency
1.	және	344
2.	мен	325
3.	бұл	308
4.	да	250
5.	деп	246
6.	астана	200
7.	бір	198
8.	baq.kz	186
9.	үшін	179
10.	бойынша	175
11.	де	167
12.	қазақ	167
13.	тамыз	161
14.	Қазақстан	156
15.	осы	154
16.	ал	143
17.	бар	141
18.	ол	136
19.	деген	134
20.	оның	128

#	lexeme
1.	и
2.	в (во)
3.	не
4.	на
5.	я
6.	он
7.	что
8.	с (со)
9.	это
10.	быть
11.	а
12.	весь (вся, все, все)
13.	они
14.	она
15.	как
16.	мы
17.	к (ко)
18.	у
19.	вы
20.	этот (эта, это, эти)

First table shows frequency list of the most 20 high ranging lexical units as a part of the frequency dictionary created by materials of modern Kazakh mass media for 2012-2013 (Kuryshzhan A.A., Omirzakova A.K., 2013) and second one demonstrates first 20 units of the most often used Russian language words for Beginners. Almost of these 20 units of both tables are ancillary and auxiliary words in Kazakh or in the Russian. Same situation is for English frequency dictionary (COCA, 2012).

Similar comparison has been conducted for frequency lists of dialogues in the belles-lettres literature of various periods, mass communication media language and TV, direct oral and mediated written interviews, talk-shows and aitys – oral traditional Kazakh poetries competitions (Boranbaev S., Danaev N., Zhumabaev A., Kuryshzhanova A., Omirzakova A., 2013). Findings of such frequency lists and developed methodologies are anticipated to be made accessible at the end of the research. However, findings received so far indicate new application opportunities to address educational objectives: thus, identifying frequency vocabulary provides for the opportunity to distribute the content and differentiate educational levels, identifying nuclear high-frequency vocabulary and less frequently used vocabulary for the professional or specialized education, calculating parameters for accurate and objective evaluation and control.

An adequate testing system based on selected frequent grammar minimums can help to organize the following steps of correction and improving teaching process more properly and exactly: to correct the lexis or grammar, to pay attention to words forms or to models of sentences. Use of frequency vocabulary and grammar minimums at particular level or specialty allows narrowing and detailing the studied content till actually necessary within the limited courses. A developed program and research-based toolkit provides for the opportunity to address multiple specific objectives of Kazakh educators, methodologists and teachers.

Even the preliminary results of this research show possible use intended to solve educational objectives: thus, selected frequency-specific vocabulary helps teaching content distribution, specify the language levels differentiation, identifying nuclear high-frequency vocabulary for basic, intermediate or advanced courses, devise general and professional or specialized education, and clearly calculates parameters for accurate assessment and next correction of found difficulties. Use of frequent vocabulary and grammar minimums for a particular level or specialty allows reducing an amount of necessary materials and detailing the studied content until the one actually needed within the limited courses. Software-research toolkit developed provides possibilities to solve multiple objectives of a Kazakh linguistic theory, applied language products and teaching practice of Kazakh.

References

- Boranbaev S., Danaev N., Zhumabaev A., Kuryshzhanova A., Omirzakova A. Kazakh tili koldanbaly lingvistikasynyn zhijilik zertteu maseleleri / The issues of frequency research of applied Kazakh linguistics. (2013). Proceedings from “XLIII scientific-methodical conference (ISBN 978-601-247-843-3, Vol.2, p.100-104). Kazakhstan, Almaty: KazNU.
- Kuryshzhan A.A., Omirzakova A.K. Kazak tildi BAQ vocabuljarynyn zhijilik taldauy turaly / About the frequency analyze of Kazakh mass media. (2013). Proceedings from “Actual problems of State language policy” International scientific-methodical conference (ISBN 979-601-7344-02-8, Vol.1, pp. 262-266). Kazakhstan, Almaty: Educational-Methodical Center “Til” of Almaty Department of Languages Committee of the Ministry of Culture of Republic of Kazakhstan.
- N. J. Brown (2003) Russian Learners’ dictionary. 10000 words in frequency order. London-New York: Routledge.
- Portal of the Official Language, Committee on Languages development and public-political works of the Ministry of culture and sport of the Republic of Kazakhstan <http://til.gov.kz/wps/portal>. Lexical search.
- Dictionary of Abai Kunanbaev. Frequency dictionary (2014). Online Web site: <http://goo.gl/NShty1>.
- The Corpus of Contemporary American English (COCA) 450 mln words, 1990-2012. Online Web site: <http://corpus.byu.edu/coca/>.